Prediction of total population in Togo using ARIMA models

THABANI NYONI and CHIPO MUTONGI

University of Zimbabwe, Midlands State University

7 May 2019

PREDICTION OF TOTAL POPUALTION IN TOGO USING ARIMA MODELS

Nyoni, Thabani (MSc. Econ. Student)
Department of Economics
University of Zimbabwe
Harare, Zimbabwe
Email: nyonithabani35@gmail.com

Dr Chipo, Mutongi (Ph.D)
Department of Business Management
Midlands State University
Gweru, Zimbabwe
mutongic@gmail.com

Abstract

Using annual time series data on total population in Togo from 1960 to 2017, we model and forecast total population over the next 3 decades using the Box – Jenkins ARIMA technique. Diagnostic tests such as the ADF tests show that Togo annual total population is neither I (1) nor I (2) but for simplicity purposes, the researcher has assumed it is I (2). Based on the AIC, the study presents the ARIMA (3, 2, 0) model as the best model. The diagnostic tests further indicate that the presented model is stable. The results of the study reveal that total population in Togo will continue to rise in the next three decades and in 2050 Togo's total population will be approximately 14.2 million people. In order to benefit from an increase in total population in Togo, 3 policy recommendations have been suggested for consider by the government of Togo.

Key Words: Forecasting, Population, Togo

JEL Codes: C53, Q56, R23

INTRODUCTION

As the 21st century began, the world's population was estimated to be almost 6.1 billion people (Tartiyus *et al*, 2015). Projections by the United Nations place the figure at more than 9.2 billion by the year 2050 before reaching a maximum of 11 billion by 2200. Over 90% of that population will inhabit the developing world (Todaro & Smith, 2006). The problem of population growth is basically not a problem of numbers but that of human welfare as it affects the provision of welfare and development. The consequences of rapidly growing population manifests heavily on species extinction, deforestation, desertification, climate change and the destruction of natural ecosystems on one hand; and unemployment, pressure on housing, transport traffic congestion, pollution and infrastructure security and stain on amenities (Dominic *et al*, 2016).

The population growth rate in Togo is high and extremely variable among regions. Average annual growth is 2.84% (and doubles every 25 years) and women constitute the majority (51.4%). Togo's population is also mobile. People migrate in search economic opportunities, with some moving from rural to urban areas and others leaving the country (IMF, 2014). In Togo, just like in any other part of the world, population modeling and forecasting is important for policy dialogue. This study endeavors to model and forecast population of Togo using the Box-Jenkins ARIMA technique.

REVIEW OF PREVIOUS STUDIES

Table 1

Author(s) / Year	Country	Period	Methodology	Major Findings
Zakria & Muhammad (2009)	Pakistan	1951 – 2007	Box-Jenkins ARIMA technique	ARIMA (1, 2, 0) is the best model for forecasting total population in Pakistan
Haque *et al* (2012)	Bangladesh	1991 – 2006	Logistic Population Model (LPM)	The LPM has the best fit for population growth in Bangladesh
Beg & Islam (2016)	Bangladesh	1965 – 2003	Autoregressive Time Trend Model (ATTM)	Downward population growth for Bangladesh for the extended period up to 2043
Ayele & Zewdie (2017)	Ethiopia	1961 – 2009	Box-Jenkins ARIMA technique	ARIMA (2, 1, 2) Model is the best model for forecasting total population in Ethiopia

MATERIALS & METHODS

ARIMA Models

ARIMA models are often considered as delivering more accurate forecasts then econometric techniques (Song *et al*, 2003b). ARIMA models outperform multivariate models in forecasting performance (du Preez & Witt, 2003). Overall performance of ARIMA models is superior to that of the naïve models and smoothing techniques (Goh & Law, 2002). ARIMA models were developed by Box and Jenkins in the 1970s and their approach of identification, estimation and diagnostics is based on the principle of parsimony (Asteriou & Hall, 2007). The general form of the ARIMA (p, d, q) can be represented by a backward shift operator as:

$$\emptyset(B)(1-B)^d POP_t = \theta(B)\mu_t \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots [1]$$

Where the autoregressive (AR) and moving average (MA) characteristic operators are:

$$\emptyset(B) = \left(1-\emptyset_1 B-\emptyset_2 B^2-\cdots-\emptyset_p B^p\right) \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots [2]$$

$$\theta(B) = \left(1-\theta_1 B-\theta_2 B^2-\cdots-\theta_q B^q\right) \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots [3]$$

and

$$(1 - B)^d POP_t = \Delta^d POP_t \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots [4]$$

Where $\emptyset$ is the parameter estimate of the autoregressive component, θ is the parameter estimate of the moving average component, Δ is the difference operator, d is the difference, B is the backshift operator and μ_t is the disturbance term.

The Box – Jenkins Methodology

The first step towards model selection is to difference the series in order to achieve stationarity. Once this process is over, the researcher will then examine the correlogram in order to decide on the appropriate orders of the AR and MA components. It is important to highlight the fact that this procedure (of choosing the AR and MA components) is biased towards the use of personal judgement because there are no clear – cut rules on how to decide on the appropriate AR and MA components. Therefore, experience plays a pivotal role in this regard. The next step is the estimation of the tentative model, after which diagnostic testing shall follow. Diagnostic checking is usually done by generating the set of residuals and testing whether they satisfy the characteristics of a white noise process. If not, there would be need for model re – specification and repetition of the same process; this time from the second stage. The process may go on and on until an appropriate model is identified (Nyoni, 2018).

Data Collection

This study is based on 58 observations of annual total population in Togo. All the data was gathered from the World Bank.

Diagnostic Tests & Model Evaluation

Stationarity Tests: Graphical Analysis

Figure 1

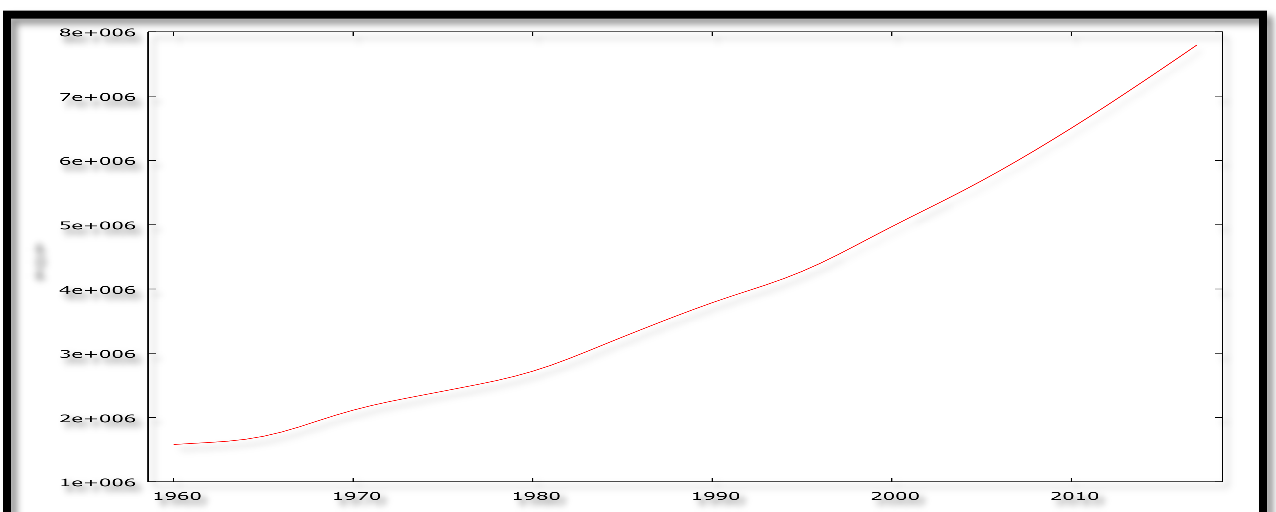
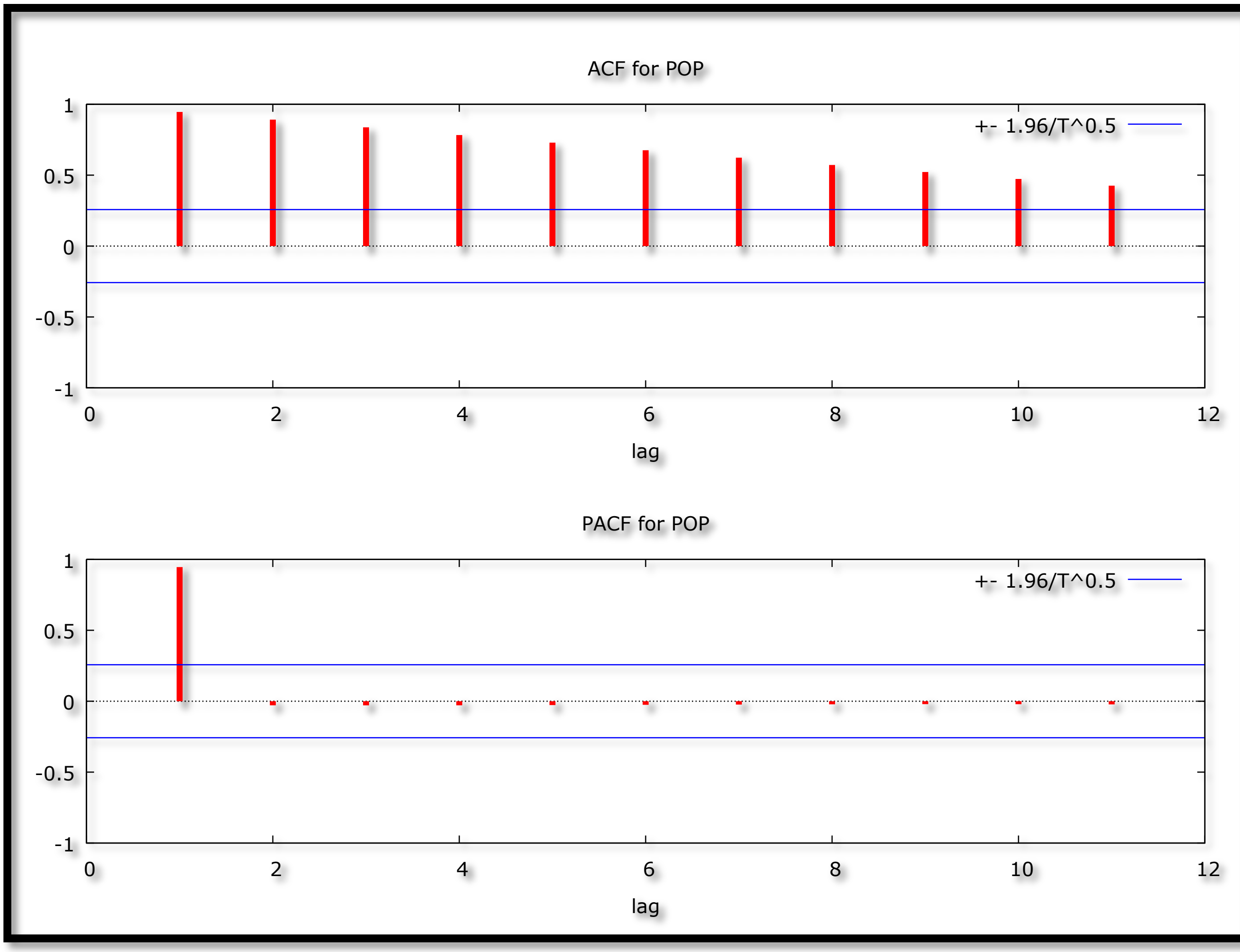


Figure 1 above indicates that the Togo POP variable is not stationary since it is trending upwards over the period 1960 – 2017. This basically means that the mean and varience of POP is changing over time.

The Correlogram in Levels

Figure 2



The ADF Test

Table 2: Levels-intercept

Variable	ADF Statistic	Probability	Critical Values		Conclusion
POP	1.240884	0.9980	-3.571310	@1%	Not stationary
			-2.922449	@5%	Not stationary
			-2.599224	@10%	Not stationary

Table 3: Levels-trend & intercept

Variable	ADF Statistic	Probability	Critical Values		Conclusion
POP	2.014416	1.0000	-4.156734	@1%	Not stationary

			-3.504330	@5%	Not stationary
			-3.181826	@10%	Not stationary

Table 4: without intercept and trend & intercept

Variable	ADF Statistic	Probability	Critical Values		Conclusion
POP	-0.093694	0.6464	-2.613010	@1%	Not stationary
			-1.947665	@5%	Not stationary
			-1.612573	@10%	Not stationary

The Correlogram (at 1st Differences)

Figure 3

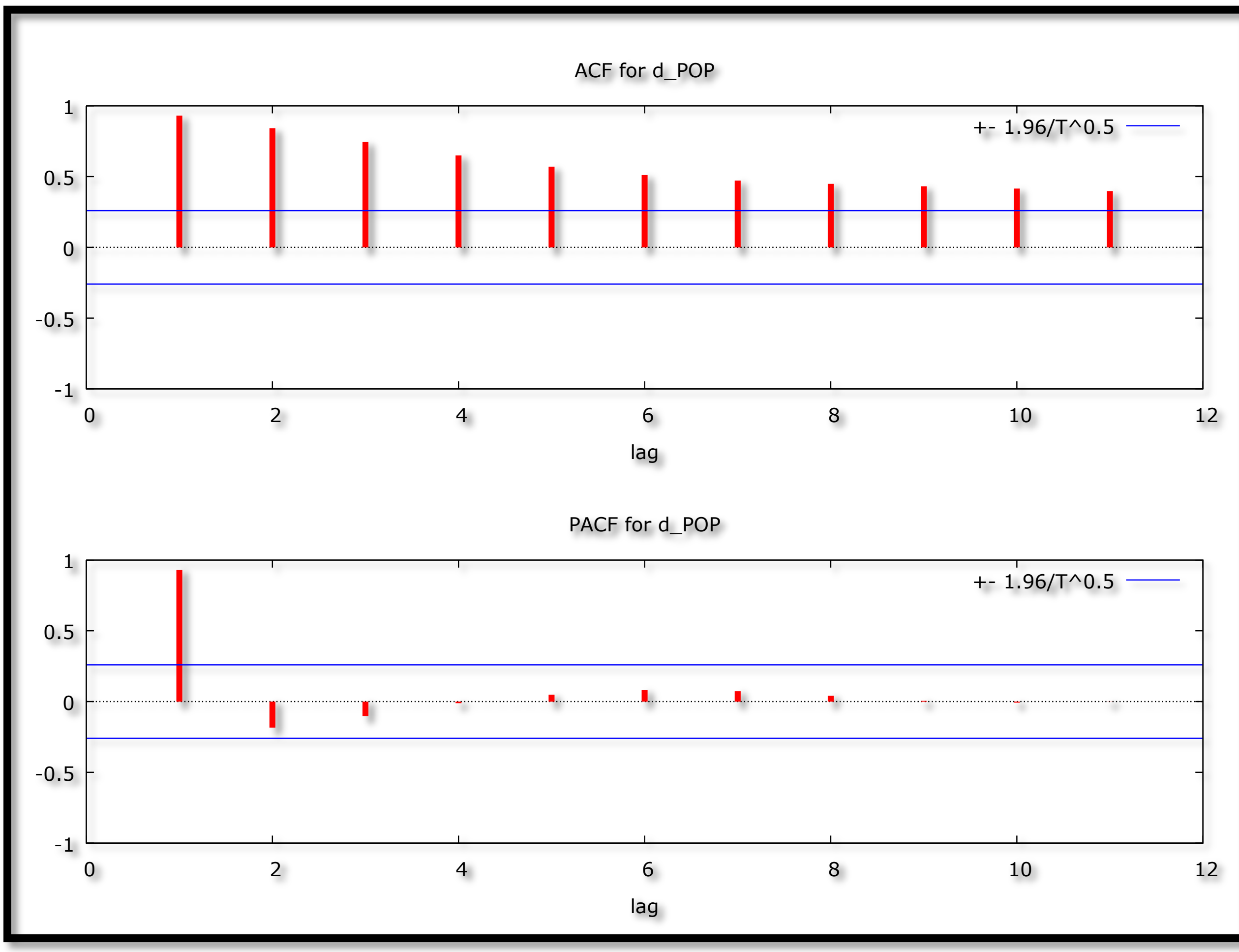


Table 5: 1st Difference-intercept

Variable	ADF Statistic	Probability	Critical Values		Conclusion
POP	0.958866	0.9955	-3.581152	@1%	Not stationary
			-2.926622	@5%	Not stationary
			-2.601424	@10%	Not stationary

Table 6: 1st Difference-trend & intercept

Variable	ADF Statistic	Probability	Critical Values		Conclusion
POP	-1.929801	0.6229	-4.170583	@1%	Not stationary
			-3.510740	@5%	Not stationary
			-3.185512	@10%	Not stationary

Table 7: 1st Difference-without intercept and trend & intercept

Variable	ADF Statistic	Probability	Critical Values		Conclusion
POP	1.874309	0.9841	-2.616203	@1%	Not stationary
			-1.948140	@5%	Not stationary
			-1.612320	@10%	Not stationary

Figures above, i.e. 2 and 3 and tables above, i.e. 2 to 7 demonstrate that the Togo POP series is not stationary in levels and in first differences.

The Correlogram in (2nd Differences)

Figure 4

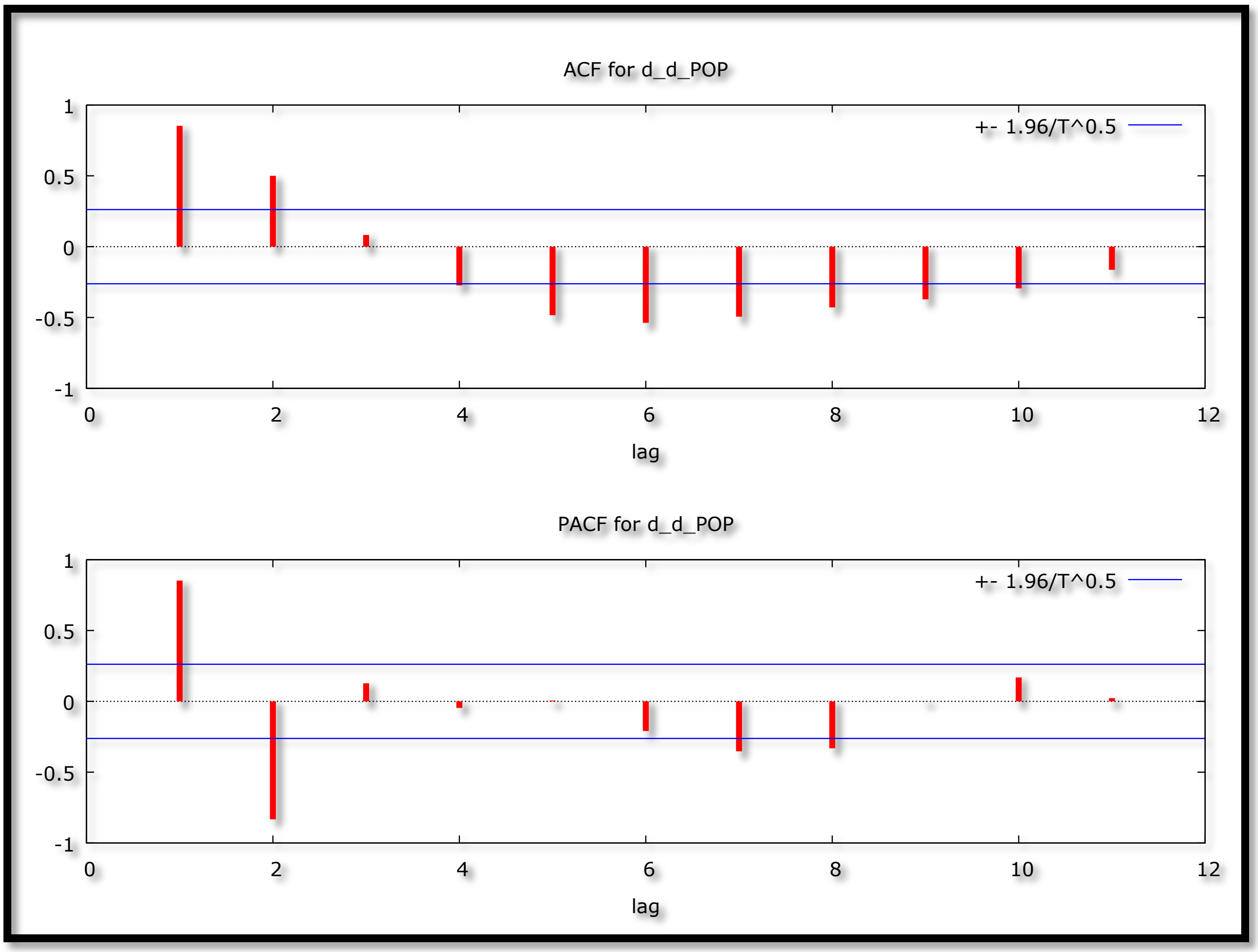


Table 8: 2nd Difference-intercept

Variable	ADF Statistic	Probability	Critical Values	Conclusion

POP	-2.040910	0.2690	-3.581152	@1%	Not stationary
			-2.926622	@5%	Not stationary
			-2.601424	@10%	Not stationary

Table 9: 2nd Difference-trend & intercept

Variable	ADF Statistic	Probability	Critical Values		Conclusion
POP	-2.368235	0.3906	-4.170583	@1%	Not stationary
			-3.510740	@5%	Not stationary
			-3.185512	@10%	Not stationary

Table 10: 2nd Difference-without intercept and trend & intercept

Variable	ADF Statistic	Probability	Critical Values		Conclusion
POP	-0.372457	0.5448	-2.616203	@1%	Not stationary
			-1.948140	@5%	Not stationary
			-1.612320	@10%	Not stationary

Figure 4 and tables 8 – 10 illustrate that the Togo POP series is not stationary even after taking second differences. This is a characteristic of sharply upwards trending series. However, the researcher will assume that the Togo POP series is I (2).

Evaluation of ARIMA models (without a constant)

Table 11

Model	AIC	U	ME	MAE	RMSE	MAPE
ARIMA (1, 2, 1)	1025.393	0.034114	296.64	1650.7	2202.6	0.061962
ARIMA (1, 2, 0)	1073.964	0.051317	427.41	2461.2	3372.8	0.092581
ARIMA (0, 2, 1)	1090.902	0.062059	1638	3096	4042.8	0.11286
ARIMA (2, 2, 1)	979.2469	0.023799	481.96	1177.1	1525.6	0.043431
ARIMA (3, 2, 1)	978.8015	0.023743	388.48	1138.7	1506.9	0.04239
ARIMA (4, 2, 1)	980.7991	0.023742	390.37	1139.6	1506.9	0.042415
ARIMA (5, 2, 1)	982.7111	0.023773	375.35	1138.7	1506.3	0.042448
ARIMA (6, 2, 1)	982.6764	0.0236	470.62	1117.2	1488.9	0.041174
ARIMA (2, 2, 0)	990.4210	0.024768	645.71	1287.5	1667.4	0.046315
ARIMA (3, 2, 0)	**977.6460**	0.023723	360.62	1115.8	1515	0.041717
ARIMA (4, 2, 0)	978.9158	0.023716	397.06	1137.7	1507.9	0.042318
ARIMA (5, 2, 0)	980.7509	0.023767	375.67	1139.3	1506.6	0.042465
ARIMA (6, 2, 0)	982.0784	0.023746	412.2	1128.2	1500.8	0.041883

A model with a lower AIC value is better than the one with a higher AIC value (Nyoni, 2018). Theil’s U must lie between 0 and 1, of which the closer it is to 0, the better the forecast method (Nyoni, 2018). The study will consider the minimum AIC in order to choose the best model for forecasting total population in Togo. Therefore, the ARIMA (3, 2, 0) model is carefully selected.

Residual & Stability Tests

ADF Tests of the Residuals of the ARIMA (3, 2, 0) Model

Table 12: Levels-intercept

Variable	ADF Statistic	Probability	Critical Values		Conclusion
R_t	-10.04156	0.0000	-3.592462	@1%	Stationary
			-2.931404	@5%	Stationary
			-2.603944	@10%	Stationary

Table 13: Levels-trend & intercept

Variable	ADF Statistic	Probability	Critical Values		Conclusion
R_t	-10.12991	0.0000	-4.186481	@1%	Stationary
			-3.518090	@5%	Stationary
			-3.189732	@10%	Stationary

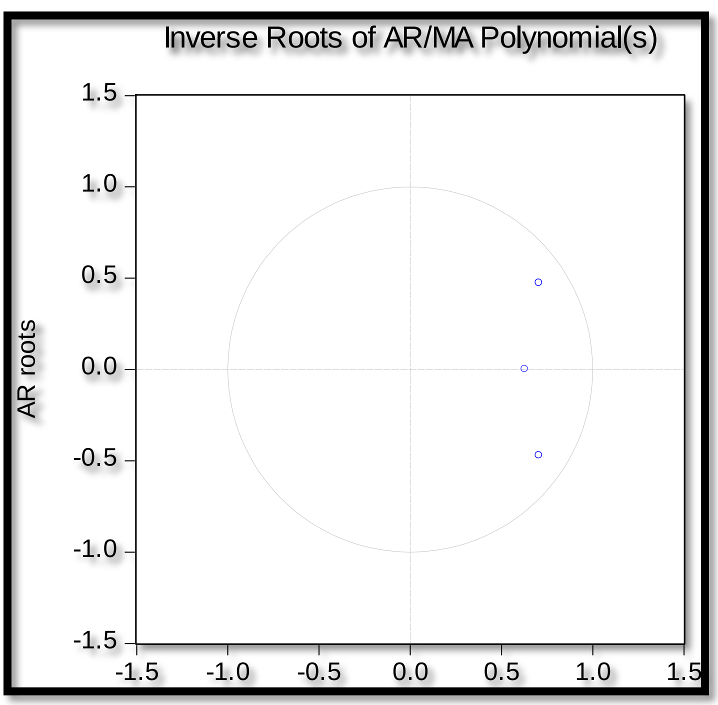
Table 14: without intercept and trend & intercept

Variable	ADF Statistic	Probability	Critical Values		Conclusion
R_t	-1.238305	0.1946	-2.621185	@1%	Not stationary
			-1.948886	@5%	Not stationary
			-1.611932	@10%	Not stationary

Tables 12 and 13 indicate that the residuals of the ARIMA (3, 2, 0) model are stationary while table 14 shows that they are not. However, we give priority to results from tables 12 and 13 since the data under consideration has a trend (it is has an upward trend as shown in figure 1 above).

Stability Test of the ARIMA (3, 2, 0) Model

Figure 5



Since the corresponding inverse roots of the characteristic polynomial lie in the unit circle, it illustrates that the chosen ARIMA (3, 2, 0) model is quite stable.

FINDINGS

Descriptive Statistics

Table 15

Description	Statistic
Mean	3951500
Median	3634200
Minimum	1580500
Maximum	7797700
Standard deviation	1852000
Skewness	0.48886
Excess kurtosis	-0.95673

As shown above, the mean is positive, i.e. 3951500. The wide gap between the minimum (i.e 1580500) and the maximum (i.e. 7797700) is consistent with the observation that the Togo POP series is gradually trending upwards over the period 1960 – 2017. The skewness is 0.48886 and the most essential characteristic is that it is positive, meaning that the Togo POP series is positively skewed and non-symmetric. Excess kurtosis is -095673; showing that the Togo POP series is not normally distributed.

Results Presentation[1]

Table 16

ARIMA (3, 2, 0) Model:

$$\Delta^2 POP_{t-1} = 2.10605\Delta^2 POP_{t-1} - 1.71697\Delta^2 POP_{t-2} + 0.503511\Delta^2 POP_{t-3} \ldots\ldots\ldots\ldots\ldots\ldots [5]$$

P: (0.0000) (0.0000) (0.0000)

S. E: (0.119144) (0.204895) (0.116762)

Variable	Coefficient	Standard Error	z	p-value
AR (1)	2.10605	0.119144	17.68	0.0000***
AR (2)	-1.71697	0.204895	-8.38	0.0000***
AR (3)	0.503511	0.116762	4.312	0.0000***

Table 17

Year	Actual POP	Fitted	Residual
1962	1612755.00	1614539.00	-1784.00
1963	1631764.00	1626405.35	5358.65
1964	1662073.00	1658628.95	3444.05

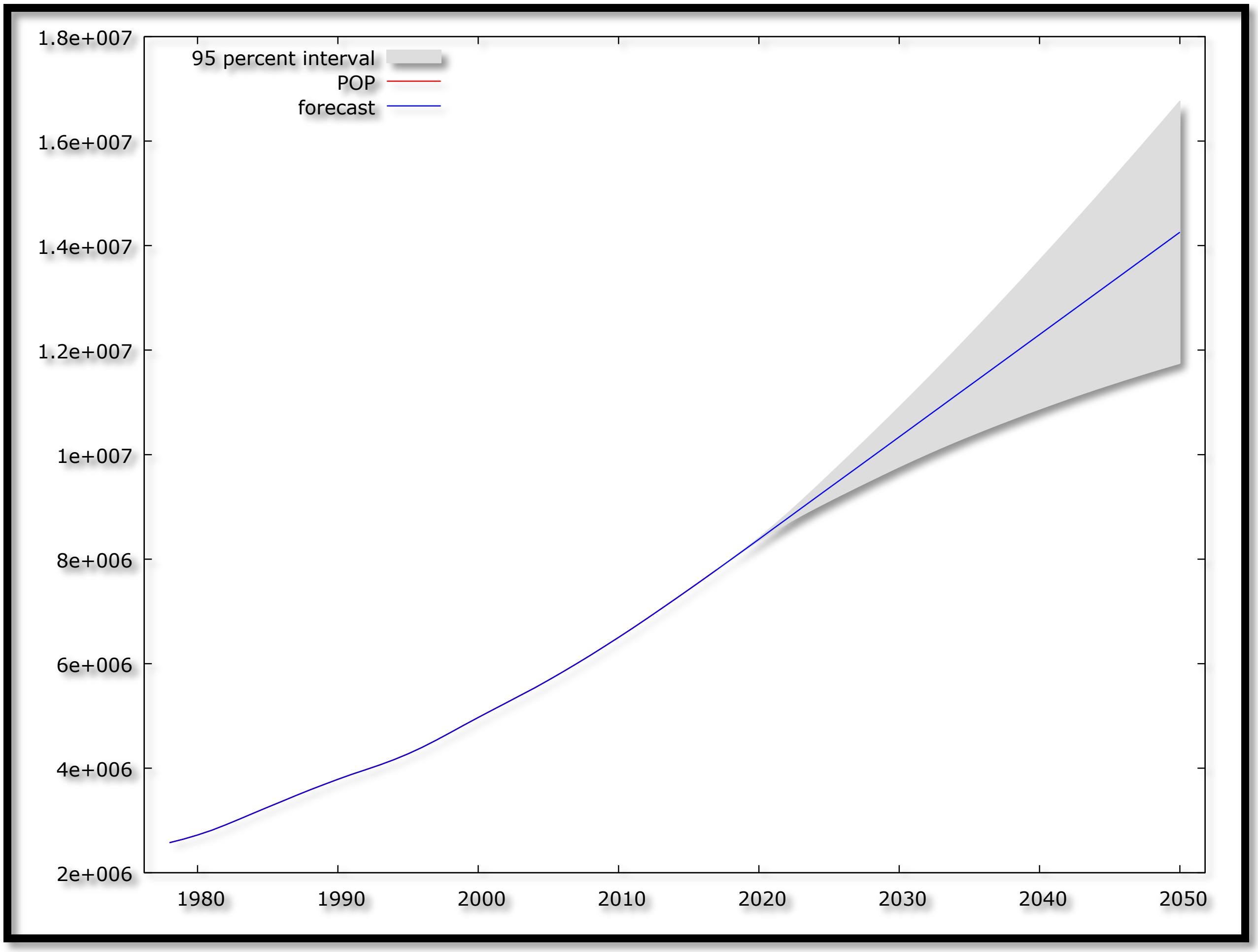
[1] *The *, ** and *** means significant at 10%, 5% and 1% levels of significance; respectively.*

1965	1708630.00	1708791.90	-161.90
1966	1774029.00	1771907.52	2121.48
1967	1855442.00	1856902.43	-1460.43
1968	1945780.00	1946411.08	-631.08
1969	2034907.00	2036906.03	-1999.03
1970	2115522.00	2114222.84	1299.16
1971	2185662.00	2184783.43	878.57
1972	2247582.00	2247746.29	-164.29
1973	2303345.00	2305889.70	-2544.70
1974	2356622.00	2354980.31	1641.69
1975	2410446.00	2411095.90	-649.90
1976	2464455.00	2466590.28	-2135.28
1977	2518566.00	2516662.71	1903.29
1978	2576469.00	2572849.60	3619.40
1979	2642846.00	2642276.14	569.86
1980	2720839.00	2720610.23	228.77
1981	2812039.00	2810655.52	1383.48
1982	2915066.00	2915375.95	-309.95
1983	3026238.00	3026173.94	64.06
1984	3140237.00	3140906.99	-669.99
1985	3252994.00	3252160.08	833.92
1986	3364020.00	3362382.51	1637.49
1987	3474080.00	3474956.34	-876.34
1988	3581928.00	3584452.28	-2524.28
1989	3686373.00	3685904.44	468.56
1990	3786940.00	3786962.68	-22.68
1991	3882271.00	3884068.84	-1797.84

1992	3973327.00	3971519.71	1807.29
1993	4064926.00	4062417.10	2508.90
1994	4163642.00	4162372.25	1269.75
1995	4274024.00	4274261.90	-237.90
1996	4398238.00	4397028.84	1209.16
1997	4534551.00	4535136.12	-585.12
1998	4679023.00	4678469.85	553.15
1999	4825704.00	4826869.15	-1165.15
2000	4970367.00	4969120.47	1246.53
2001	5111770.00	5111095.36	674.64
2002	5251472.00	5250884.40	587.60
2003	5391401.00	5392172.86	-771.86
2004	5534598.00	5533087.19	1510.81
2005	5683268.00	5683431.33	-163.33
2006	5837792.00	5837967.62	-175.62
2007	5997385.00	5996893.28	491.72
2008	6161796.00	6160358.11	1437.89
2009	6330472.00	6330598.15	-126.15
2010	6502952.00	6502410.21	541.79
2011	6679282.00	6678546.43	735.57
2012	6859482.00	6859336.39	145.61
2013	7042948.00	7043137.41	-189.41
2014	7228915.00	7228586.18	328.82
2015	7416802.00	7416490.18	311.82
2016	7606374.00	7606082.93	291.07
2017	7797694.00	7797457.38	236.62

Forecast Graph

Figure 6



Predicted Total Population

Table 18

Year	Prediction	Std. Error	95% Confidence Interval
2018	7990769.01	1326.919	7988168.30 - 7993369.73
2019	8185387.31	5607.645	8174396.53 - 8196378.09
2020	8381122.66	14320.610	8353054.78 - 8409190.54
2021	8577444.48	28268.326	8522039.58 - 8632849.39
2022	8773860.54	47427.901	8680903.56 - 8866817.52
2023	8970030.56	71088.007	8830700.63 - 9109360.49
2024	9165815.91	98164.402	8973417.21 - 9358214.60
2025	9361261.00	127549.700	9111268.19 - 9611253.82

2026	9556526.12	158381.266	9246104.54 - 9866947.69
2027	9751802.67	190170.938	9379074.48 - 10124530.86
2028	9947241.05	222800.461	9510560.17 - 10383921.93
2029	10142909.95	256424.473	9640327.21 - 10645492.68
2030	10338792.26	291333.106	9767789.86 - 10909794.65
2031	10534809.71	327818.630	9892297.00 - 11177322.42
2032	10730861.41	366077.193	10013363.29 - 11448359.52
2033	10926860.67	406162.596	10130796.61 - 11722924.73
2034	11122758.73	447993.451	10244707.70 - 12000809.76
2035	11318550.93	491399.704	10355425.21 - 12281676.65
2036	11514267.56	536184.894	10463364.48 - 12565170.64
2037	11709955.81	582180.718	10568902.57 - 12851009.06
2038	11905660.77	629278.938	10672296.72 - 13139024.83
2039	12101411.58	677437.056	10773659.35 - 13429163.81
2040	12297215.97	726663.381	10872981.91 - 13721450.03
2041	12493062.91	776991.685	10970187.19 - 14015938.63
2042	12688930.53	828455.781	11065187.04 - 14312674.03
2043	12884795.66	881071.531	11157927.19 - 14611664.12
2044	13080641.41	934829.611	11248409.04 - 14912873.78
2045	13276461.08	989698.347	11336687.96 - 15216234.19
2046	13472257.83	1045633.065	11422854.68 - 15521660.97
2047	13668041.31	1102587.350	11507009.82 - 15829072.81
2048	13863823.11	1160522.198	11589241.40 - 16138404.82
2049	14059612.56	1219410.743	11669611.42 - 16449613.70
2050	14255414.37	1279238.260	11748153.45 - 16762675.29

Table 17 shows the actual total population of Togo, the fitted one as well as the residuals. The striking feature of table 17 is the residuals are reasonably small, confirming the accuracy of the selected model, the ARIMA (3, 2, 0) model as already hinted by the forecast evaluation statistics in table 11 above. Figure 6 (with a forecast range from 2018 – 2050) and table 18, clearly show that Togo's total population is set to continue rising gradually, in the next 3 decades. With a 95% confidence interval of 11748153 to 16762675 and a projected total population of 14255414 by 2050, the chosen ARIMA (3, 2, 0) model is consistent with the population projections by the UN (2015) which forecasted that Togo's population will be approximately 15681000 by 2050.

Policy Implications

i. The government of Togo ought to invest more in infrastructural development in order to cater for the expected increase in total population.
ii. The predicted increase in total population justifies the need for more and bigger companies to provide for the anticipated increase in demand for goods and services in Togo.
iii. The government of Togo should take action so as to improve health service delivery in the country in order to ensure a healthier society, particularly in light of such a likely increase in total population.

CONCLUSION

In the case of Togo, the study shows that the ARIMA (3, 2, 0) model is not only stable but also the most suitable model to forecast total population for the next 3 decades. The model predicts that by 2050, Togo's total population would be approximately, 14.2 million people. This is a warning signal to policy makers in Togo, particularly with regards to infrastructural development, e.g schools and hospitals. These findings are essential for the government of Togo, especially when it comes to long-term planning.

REFERENCES

[1] Asteriou, D. & Hall, S. G. (2007). Applied Econometrics: a modern approach, Revised Edition, *Palgrave MacMillan*, New York.

[2] Ayele, A. W & Zewdie, M. A (2017). Modeling and forecasting Ethiopian human population size and its pattern, *International Journal of Social Sciences, Arts and Humanities*, 4 (3): 71 – 82.

[3] Beg, A. B. M. R. A & Islam, M. R (2016). Forecasting and modeling population growth of Bangladesh, *American Journal of Mathematics and Statistics*, 6 (4): 190 – 195.

[4] Dominic, A., Oluwatoyin, M. A., & Fagbeminiyi, F. F (2016). The determinants of population growth in Nigeria: a co-integration approach, *The International Journal of Humanities and Social Studies*, 4 (11): 38 – 44.

[5] Du Preez, J. & Witt, S. F. (2003). Univariate and multivariate time series forecasting: An application to tourism demand, *International Journal of Forecasting*, 19: 435 – 451.

[6] Goh, C. & Law, R. (2002). Modeling and forecasting tourism demand for arrivals with stochastic non-stationary seasonality and intervention, *Tourism Management*, 23: 499 – 510.

[7] Haque, M., Ahmed, F., Anam, S., & Kabir, R (2012). Future population projection of Bangladesh by growth rate modeling using logistic population model, *Annals of Pure and Applied Mathematics*, 1 (2): 192 – 202.

[8] IMF (2014). Togo poverty reduction strategy paper, IMF country report, No. 14/224, *IMF*, Washington DC.

[9] Nyoni, T (2018). Modeling and Forecasting Naira / USD Exchange Rate in Nigeria: a Box – Jenkins ARIMA approach, *University of Munich Library – Munich Personal RePEc Archive (MPRA)*, Paper No. 88622.

[10] Nyoni, T (2018). Modeling and Forecasting Inflation in Kenya: Recent Insights from ARIMA and GARCH analysis, *Dimorian Review*, 5 (6): 16 – 40.

[11] Nyoni, T. (2018). Box – Jenkins ARIMA Approach to Predicting net FDI inflows in Zimbabwe, *Munich University Library – Munich Personal RePEc Archive (MPRA)*, Paper No. 87737.

[12] Song, H., Witt, S. F. & Jensen, T. C. (2003b). Tourism forecasting: accuracy of alternative econometric models, *International Journal of Forecasting*, 19: 123 – 141.

[13] Tartiyus, E. H., Dauda, T. M., & Peter, A (2015). Impact of population growth on economic growth in Nigeria, *IOSR Journal of Humanities and Social Science (IOSR-JHSS)*, 20 (4): 115 – 123.

[14] Todaro, M & Smith, S (2006). Economic Development, 9th Edition, *Vrinda* Publications, New Delhi.

[15] United Nations (2015). World Population Prospects: The 2015 Revision, Key Findings and Advance Tables, *Department of Economic and Social Affairs*, Population Division, Working Paper No. ESA/P/WP/241.

[16] Zakria, M & Muhammad, F (2009). Forecasting the population of Pakistan using ARIMA models, *Pakistan Journal of Agricultural Sciences*, 46 (3): 214 – 223.